Philosophy's Gender Gap and Argumentative Arena: An Empirical Study

Moti Mizrahi, Florida Institute of Technology

Mike Dickinson, University of Illinois

Forthcoming in *Synthese*

Abstract: While the empirical evidence pointing to a gender gap in professional, academic philosophy in the English-speaking world is widely accepted, explanations of this gap are less so. In this paper, we aim to make a modest contribution to the literature on the gender gap in academic philosophy by taking a quantitative, corpus-based empirical approach. Since some philosophers have suggested that it may be the argumentative, "logic-chopping," and "paradox-mongering" nature of academic philosophy that explains the underrepresentation of women in the discipline, our research questions are the following: Do men and women philosophers make different types of arguments in their published works? If so, which ones and with what frequency? Using data mining and text analysis methods, we study a large corpus of philosophical texts mined from the JSTOR database in order to answer these questions empirically. Using indicator words to classify arguments by type (namely, deductive, inductive, and abductive arguments), we search through our corpus to find patterns of argumentation. Overall, the results of our empirical study suggest that women philosophers make deductive, inductive, and abductive arguments in their published works just as much as male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work.

Keywords: argumentation; corpus linguistics; data science; gender gap; metaphilosophy; philosophical methodology; text mining; women in philosophy

1. Introduction

Anyone who is familiar with professional philosophy, i.e., the academic discipline as it is practiced in the English-speaking world, knows that it has a "gender problem" or, more specifically, a "women problem." As Bowell (2015, p. 4) puts it, "Philosophy has always had, and continues to have, a Woman Problem. Women remain *under-represented* as students, as scholars, in journal publications – especially in 'top' journals – and as philosophical subjects" (emphasis added).[1] Accordingly, the "women problem" and "gender problem" labels refer to a complex "set of problems that relate to the general underrepresentation of women in philosophy: in the historical canon, in the professoriate class, at conferences [Schwitzgebel and Jennings (2017)], and in upper-level undergraduate classrooms [Paxton et al. (2012)]" (Krishnamurthy et al. 2017, p. 928).

While the empirical evidence pointing to a gender gap in academic philosophy is widely accepted, explanations of this gap are less so.[2] According to Friedman (2013, p. 25), the "likeliest aspects of philosophy that might deter or alienate women are (1) the contents of philosophy, its distinctive questions, issues, and ideas; (2) the *methods*, broadly construed, that are used in philosophy to deal with that content; and (3) the way in which philosophy is taught and communicated" (emphasis added). With respect to (2), in particular, Friedman (2013, p. 25) cites Toulmin as describing "a philosophical method or approach that could alienate women from the field." On this philosophical method or approach, which Toulmin describes but does not endorse, academic philosophers are characterized as "just logic-choppers and paradox-mongers

[1] For data on the representation of women philosophers in philosophy journals, see Wilhelm et al. (2018).
[2] For an overview of explanations of the gender gap in academic philosophy, see Thompson (2017). See also Thompson et al. (2016) and Easton (2021).

who trade on the confusions produced by playing word games with tricky abstract nouns" (Toulmin 1976, p. 11). "If this is philosophy," Friedman (2013, p. 25) argues, then it is "no wonder women have relatively little interest in entering the field." Indeed, Friedman (2013, p. 25) goes on to argue, "It is not clear why men would enter the field either, unless men were more enamored than women of 'logic-chopping' and 'paradox-mongering'."

Along the same lines, Demarest et al. (2017, p. 530) propose that "When an instructor explicitly presents a counter-stereotypical profile of an author, it directly combats the existing stereotype of a philosopher as a white man (without family, obligations, or other interests and *concerned only with arguments*)" (emphasis added). This proposal to counter the stereotype of an academic philosopher who is concerned only with arguments is motivated by their empirical findings, which suggest that "two attitudes are especially strong predictors of whether women are likely to continue in philosophy: (i) feeling similar to the kinds of people who become philosophers and (ii) enjoying philosophical puzzles and issues" (Demarest et al. 2017, p. 526).[3]

Even if we should do something to combat the existing stereotype of a male and white academic philosopher who is concerned only with arguments, as Demarest et al. (2017, p. 530) recommend, one would be hard pressed to deny that academic philosophers are in fact concerned with arguments. After all, to do philosophy professionally is to put forth arguments for and against philosophical theses. And logic is the study of arguments. So, logic and argumentation are indispensable to doing philosophy professionally. As Harrell (2016, p. 7) observes, "philosophers are primarily concerned with arguments." But if logic and argumentation are indispensable to doing philosophy professionally, and there is a gender gap in academic philosophy, the following question arises: could this gap have something to do with the ways in

[3] For more on what can be done to improve the climate for women in academic philosophy, see Wuest (2013).

which men and women philosophers argue? In other words, if the methods of philosophy, which include philosophical argumentation, are one of the aspects of philosophy that might deter or alienate women, as Friedman (2013, p. 25) argues, it seems reasonable to think that the gender gap in academic philosophy might have something to do with differences in philosophical argumentation between men and women philosophers.[4] By contrast, Warnock (2015) claims that academic philosophy's "women problem" cannot "be explained by the supposition that, philosophy being concerned above all with arguments, women are naturally less adept in the field."

In this paper, we take a quantitative, corpus-based empirical approach to this question, thereby aiming to make a modest contribution to the literature on the gender gap in academic philosophy. More explicitly, the following empirical questions are the research questions that guide our empirical study in this paper:

1. Do men and women philosophers make different types of arguments in their published works?
2. If men and women philosophers make different types of arguments in their published works, which types of arguments are typically made by male philosophers and which types of arguments are typically made by women philosophers? Are there significant differences between the types of arguments typically made by male philosophers versus those typically made by women philosophers?

[4] Another prevalent method that is used in academic philosophy to deal with philosophical content is eliciting intuitions through thought experiments. Buckwalter and Stich (2014) present empirical evidence suggesting that men and women have different philosophical intuitions. Adleberg et al. (2015) tried to replicate the findings reported by Buckwalter and Stich (2014) but did not find any evidence that men and women have different intuitions about philosophical thought experiments. See also Antony (2012).

We set out to investigate these questions empirically. In other words, we set out to test "the supposition that, philosophy being concerned above all with arguments, women are naturally less adept in the field" (Warnock 2015) empirically. Using data mining and text analysis methods, we study a large corpus of philosophical texts mined from the JSTOR database. Using indicator words to classify arguments by type (namely, deductive, inductive, and abductive arguments), we search through our corpus to find patterns of argumentation. Before we report the results of our empirical study in Section 3, we describe our methodology in more detail in Section 2. In Section 4, we will discuss how the results of our empirical study provide tentative answers to our research questions (1) and (2) above. Overall, the results of our empirical study suggest that women philosophers make deductive, inductive, and abductive arguments in their published works just as much as male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work.

From the outset, then, it is important to note that our empirical study focuses on arguments made in academic articles published in philosophy journals. Of course, arguments made in academic publications are an integral part of the argumentative landscape of academic philosophy. But there are other aspects to argumentation in academic philosophy that go beyond the journal article. In addition to making arguments in academic publications, academic philosophers also make arguments in the seminar room and while giving talks at academic conferences. For example, Alcoff (2013) recounts the following experience: "Chisholm was typical of the best philosophers of his day and ours in his combination of philosophical acumen and rhetorical skill. Yet he was atypical at that time in his sensitivity to the practical contexts of *the argumentative arena*. [...]. As one of two women in the class, he was aware I might be experiencing an alienation-induced anxiety about my public performance" (emphasis added).

These aspects of academic philosophy's argumentative landscape or "argumentative arena" (Alcoff 2013), such as the philosophy seminar (Beebee 2013, pp. 63-73) and the academic conference, are not the focus of our empirical study. Rather, our focus is on arguments in the more formal sense of premises given in support of a conclusion, not arguments in the more informal sense of a debate or a disagreement.[5] In other words, the focus of this empirical study is the *scholarly* aspect of academic philosophy's argumentative landscape or "argumentative arena" (Alcoff 2013).

2. Methods

2.1 *Background*

Introductory textbooks to logic, philosophy, and argumentation typically contain a brief discussion of indicator words. For "Indicator words suggest the presence of an argument and help to indicate its structure" (Govier 2013, p. 4). There are premise indicators, words such as 'because' and phrases such as 'inferred from' and the like, which indicate a premise of an argument, and there are conclusion indicators, words such as 'therefore' and phrases such as 'it follows that' and the like, which indicate a conclusion of an argument. For example, Morrow and Weston (2019, p. 5) instruct students to look for indicator words in order to distinguish between premises and conclusions of arguments. According to Morrow and Weston (2019, p. 5):

[5] Of course, arguments made in academic publications can be adversarial and aggressive as well. For example, in his review of Ted Honderich's *On Consciousness* (2004), McGinn (2007, p. 474) writes, "This book runs the full gamut from the mediocre to the ludicrous to merely bad. It is painful to read, poorly thought out, and uninformed. It is also radically inconsistent. [...] The second half tries to develop a new theory of consciousness, according to which the positive theses of the first half of the book are all wrong [...], and the fact is only slyly acknowledged toward the end of the discussion--hence the radical inconsistency I mentioned."

> Some words or phrases are *conclusion indicators*. These are words or phrases that tell you that you're about to read or hear the conclusion of an argument. Other words or phrases are *premise indicators*. These tell you that you're about to read or hear a premise (emphasis in original).

Morrow and Weston (2019, p. 5) then provide a list of premise indicators, which includes words like ‘because’ and phrases like ‘this follows from’, and a list of conclusion indicators, which includes words like ‘therefore’ and phrases like ‘this shows that’ (Morrow and Weston 2019, p. 5). Likewise, according to Marcus (2018, pp. 9-10), “there are premise and conclusion indicators. ‘We may conclude that’ is used to indicate a conclusion. ‘This may be inferred from the fact that’ is used to indicate a premise.”

In addition to helping students identify premises and conclusions of arguments, indicators also help students distinguish between deductive arguments and inductive arguments. For example, according to Baronett (2016, p. 23):

> To help identify arguments as either deductive or inductive, one thing we can do is look for key words or phrases. For example, the words “necessarily,” “certainly,” “definitely,” and “absolutely” suggest a deductive argument. [...] On the other hand, the words “probably,” “likely,” “unlikely,” “improbable,” “plausible,” and “implausible” suggest inductive arguments.

Similarly, according to Hurley and Watson (2018, p. 35), “inductive indicators” include terms and phrases such as ‘probably’, ‘improbable’, ‘plausible’, ‘implausible’, ‘likely’, ‘unlikely’, and ‘reasonable to conclude’, whereas “deductive indicators” include terms and phrases such as ‘it necessarily follows that’, ‘certainly’, ‘absolutely’, and ‘definitely’. As Walton (1999, p. 29) puts

it, "as we look at each inferent step in the chain of reasoning, we can classify it as a *deductive* argument [or] an *inductive* argument [...]. Our clues to this are both the context and the *indicator words* of how the argument is being used" (emphasis added).

We can use these deductive indicators and inductive indicators, then, to look for deductive arguments and inductive arguments in philosophical texts in much the same way that students use them to identify arguments in any text. To the aforementioned deductive and inductive indicators, we can also add indicators for abductive arguments, i.e., arguments in which the conclusion is supposed to be the best explanation for some phenomenon (Govier 2013, pp. 298-302). Abductive indicators include phrases such as 'account for', 'best explain', 'make sense of', and 'best explanation for' (Overton 2013, pp. 1386-1387). The types of arguments we searched for in this empirical study and their associated indicators are listed in Table 1.[6]

Table 1. Types of arguments and their indicators with examples from philosophical texts

Argument Types	**Indicators**	**Examples**
Abductive	account for, best explain, make sense of, best explanation for	"we likewise have good reason to think that certain normative moral statements are true because those statements provide the *best explanations* of some of the (non-normative) observations we have made" (Sober 2015, p. 906).
Deductive	absolutely, certainly, definitely, necessarily	"its being one's duty to do something does not *necessarily* imply that, all things considered, one should do it. So, although I admit that the goal of a theory of duty is to draw a line below which morally decent people must not go, I

[6] Ashton and Mizrahi (2018, p. 58) use a similar methodology to test the hypothesis that "philosophy is a priori and in the business of discovering necessary truths from the armchair." See also Mizrahi and Dickinson (2021).

		seem to be suggesting that, after all, people may go below that line" (Wolf 1986, p. 145).
Inductive	likely, unlikely, probably, improbable	"It is therefore *unlikely* that there is a single complete and fixed commentary per model, stable across audiences and contexts" (Maki 2009, p. 39).

Of course, we have to keep in mind that these abductive, deductive, and inductive indicators are just that--*indicators*. They are not sure signs for the presence (or absence) of arguments in texts. In other words, "the mere occurrence of an indicator word by no means guarantees the presence of an argument" (Hurley and Watson 2018, p. 16). Nevertheless, indicator words are still useful and reliable indicators of the presence of arguments in text, which is why students of philosophy, logic, and argumentation are instructed to look for them. As Lepore and Cumming (2013, p. 6) put it, "Although there are no sure signs of whether an argument is present, fairly reliable indicators exist." Lepore and Cumming (2013, p. 6) proceed to list some of the aforementioned indicator words as those listed in Table 1. In addition, since our aim is to study arguments made by professional philosophers, which are published in academic journals of philosophy, and academic philosophers are "trained in the ways of argument" (Currie 2016, p. 200), we can be quite confident that, as professional arguers, academic philosophers rarely misuse indicators in an effort to make non-arguments appear as arguments (see also Ashton and Mizrahi 2018, p. 62).

The quantitative methods we use in this empirical study, namely, text mining and corpus analysis, allow us to overcome the limitations of relying on selective quotation. After all, one can easily find instances of the aforementioned indicator words in philosophical texts written by both

men and women philosophers (see Table 1). However, selected quotations may or may not be representative of academic philosophy as a whole. By using data mining and text analysis methods, we can study a large corpus of philosophical texts, and thus obtain a broader view of the argumentative landscape in scholarly philosophical practice. Having a broader view of the argumentative landscape, or "argumentative arena" (Alcoff 2013), of academic philosophy is important for the purpose of this empirical study, since we would like to find out what role, if any, it plays in explaining the gender gap in academic philosophy.

Of course, empirical methodologies have limitations of their own. As far as the methods of data mining and text analysis are concerned, there are two major limitations. First, we can only study and analyze what is explicitly mentioned in the corpus. For the purpose of this empirical study, then, our corpus of philosophical texts must contain explicit mentions of the indicator words listed in Table 1, for us to be able to analyze ratios, means, and patterns of usage. It is reasonable to assume that there would be such explicit mentions of the indicator words listed in Table 1 in philosophical texts if academic philosophers are indeed "trained in the ways of argument" (Currie 2016, p. 200).

Second, as with any empirical methodology, there may be some false positives and/or false negatives. When it comes to the methods of text mining and corpus analysis, false negatives could occur when we search for a specific word *w* in a corpus, but do not find it, even though the corpus contains a synonym of *w*. For example, although unlikely, it is possible that our corpus of philosophical texts contains no instances of 'probably', and so a search for 'probably' would return zero results, because academic philosophers use 'likely' instead of 'probably' in all the philosophical texts that make up our corpus. On the other hand, false positives could occur when we find instances of a word *w* in our corpus, but those instances contain irrelevant uses of *w*. For

the purpose of this empirical study, then, the corpus of philosophical texts must contain not only explicit mentions of the abductive, deductive, and inductive indicators listed in Table 1, but also explicit mentions of those indicators in the context of argumentation. For example, instances of 'certainly' that occur outside of any argumentative context would be considered false positives for the purposes of this empirical study.

Now, there are a couple of things we can do to overcome the limitations of our quantitative, corpus-based approach. First, we can refine our searches by expanding our search terms to include as many indicator words as we can. For each argument type, we have four indicator words (see Table 1). This search algorithm is designed to minimize the number of false negatives, i.e., occurrences of abductive, deductive, and inductive arguments in philosophical texts that are indicated by words other than the standard ones, such as 'best explain', 'necessarily', and 'probably', by using synonymous indicator words and phrases, such as 'account for', 'certainly', and 'likely'.

Second, we can further refine our searches by pairing the argument type indicators with indicator words for arguments, such as 'therefore' and 'hence'. Since the aim of this paper is to find out what types of arguments academic philosophers actually make in scholarly philosophical practice, we need to search for the abductive, deductive, and inductive indicators listed in Table 1 in argumentative contexts by pairing the abductive, deductive, and inductive indicators listed in Table 1 with indicators words for arguments, such as 'therefore' and 'hence'. Although they are frequently mentioned as premise indicators, we chose not to use the words 'since' and 'because'. For, as Copi et al. (2011, p. 18) point out, "those words are used both in explanations and in arguments." Instead, words like 'therefore' and 'hence' tend to indicate arguments rather than explanations more reliably. By anchoring the abductive, deductive, and inductive indicators

listed in Table 1 to argument indicators, such as ‘therefore’ and ‘hence’, we can be quite confident that our indicators for argument types (see Table 1) actually indicate arguments in the corpus, and thus that the number of false positives will be minimized. This procedure results in the argument indicator pairs listed in Table 2.

Table 2. Indicator pairs for deductive, inductive, and abductive arguments

Deductive indicator pairs	**Inductive indicator pairs**	**Abductive indicator pairs**
therefore necessarily	therefore probably	therefore account for
therefore certainly	therefore likely	therefore best explain
therefore definitely	therefore unlikely	therefore make sense of
therefore absolutely	therefore improbable	therefore best explanation for
hence necessarily	hence probably	hence account for
hence certainly	hence likely	hence best explain
hence definitely	hence unlikely	hence make sense of
hence absolutely	hence improbable	hence best explanation for
so necessarily	so probably	so account for
so certainly	so likely	so best explain
so definitely	so unlikely	so make sense of
so absolutely	so improbable	so best explanation for
consequently necessarily	consequently probably	consequently account for
consequently certainly	consequently likely	consequently best explain
consequently definitely	consequently unlikely	consequently make sense of
consequently absolutely	consequently improbable	consequently best explanation for
proves necessarily	proves probably	proves account for
proves certainly	proves likely	proves best explain

proves definitely	proves unlikely	proves make sense of
proves absolutely	proves improbable	proves best explanation for
thus necessarily	thus probably	thus account for
thus certainly	thus likely	thus best explain
thus definitely	thus unlikely	thus make sense of
thus absolutely	thus improbable	thus best explanation for
follows necessarily	follows probably	follows account for
follows certainly	follows likely	follows best explain
follows definitely	follows unlikely	follows make sense of
follows absolutely	follows improbable	follows best explanation for
accordingly necessarily	accordingly probably	accordingly account for
accordingly certainly	accordingly likely	accordingly best explain
accordingly definitely	accordingly unlikely	accordingly make sense of
accordingly absolutely	accordingly improbable	accordingly best explanation for
infer necessarily	infer probably	infer account for
infer certainly	infer likely	infer best explain
infer definitely	infer unlikely	infer make sense of
infer absolutely	infer improbable	infer best explanation for

By searching for these deductive, inductive, and abductive indicator pairs (as listed in Table 2) in our corpus, we can find out what types of arguments academic philosophers make in their published works and with what frequency. For each of the indicator pairs listed in Table 2, we ran three kinds of searches: (a) a search allowing for up to three words between argument type indicator, e.g., ‘necessarily’, and argument indicator, e.g., ‘therefore’, (b) a search allowing for up to six words between argument type indicator, e.g., ‘probably’, and argument indicator,

e.g., ‘hence’, and (c) a search allowing for up to ten words between argument type indicator, e.g., ‘account for’, and argument indicator, e.g., ‘so’. This methodology is designed to help us find answers to research questions (1) and (2) above while minimizing the number of false positives and false negatives.

In order to find answers to research questions (1) and (2) above, we need to be able to distinguish between not only types of arguments (namely, deductive, inductive, or abductive arguments) but also the philosophers who make those arguments. More specifically, we need to tag the articles in our corpus as written by either men or women philosophers in order to be able to say what types of arguments are made by male philosophers and what types of arguments are made by women philosophers. For the purposes of this empirical study, we selected the names of men and women philosophers from Eric Schwitzgebel’s (2019) list of the most cited contemporary philosophers in the *Stanford Encyclopedia of Philosophy* (SEP).[7] This procedure results in the 32 names of male philosophers and 32 names of women philosophers listed in Table 3.

Table 3. A list of 32 men and 32 women philosophers most cited in the SEP

Men	**Women**
Lewis, David K.	Nussbaum, Martha
Quine, W.V.O.	Anscombe, G.E.M.
Putnam, Hilary	Korsgaard, Christine
Rawls, John	Anderson, Elizabeth
Davidson, Donald	Thomson, Judith Jarvis
Kripke, Saul	Cartwright, Nancy
Williams, Bernard	Annas, Julia

[7] Available at http://schwitzsplinters.blogspot.com/2019/08/the-295-most-cited-contemporary-authors.html. Schwitzgebel et al. (2018) have used a longer list (of 100 most-cited recent authors in the SEP) to study what they call the “insularity of Anglophone Philosophy” empirically.

Nozick, Robert	Young, Iris Marion
Williamson, Timothy	Millikan, Ruth G.
Jackson, Frank	Foot, Philippa
Nagel, Thomas	Stump, Eleonore
Searle, John R.	Okin, Susan Moller
Van Fraassen, Bas	Butler, Judith
Armstrong, David M.	O'Neill, Onora
Dummett, Michael	Zagzebski, Linda
Fodor, Jerry	Baker, Lynne Rudder
Harman, Gilbert	Haslanger, Sally
Chisholm, Roderick	Thomasson, Amie
Dennett, Daniel C.	Hurley, Susan
Chalmers, David J.	Longino, Helen
Strawson, P. F.	MacKinnon, Catharine
Stalnaker, Robert	Marcus, Ruth Barcan
Scanlon, T. M.	Benhabib, Seyla
Dworkin, Ronald	Paul, L. A.
Pettit, Philip	Alcoff, Linda Martín
Fine, Kit	Gendler, Tamar
Sober, Elliott	Wolf, Susan
Van Inwagen, Peter	Adams, Marilyn McCord
Popper, Karl	Baier, Annette
Parfit, Derek	Kamm, Frances
Kitcher, Philip	Langton, Rae
Bennett, Jonathan	Lloyd, Elisabeth

By searching for the argument indicator pairs listed in Table 2 in the published works of the men and women philosophers listed in Table 3, we can find out what types of arguments men and women philosophers make in their published works and with what frequency.

We have taken the names of 32 male philosophers and 32 women philosophers from Schwitzgebel et al. (2018) in order to make our task more manageable, given that identifying individuals as either 'man' or 'woman' is no straightforward task. But one might worry that, in

doing so, we have introduced an element of the so-called "survivor bias" into our sample of women philosophers. For we have selected women philosophers who have managed to be successful in academic philosophy *in spite of* the "logic-chopping" and "paradox-mongering" nature of argumentation in academic philosophy. According to Beasley (2018, p. 169), the so-called "survivor bias" (or "survivorship bias") is "an error in our logic that comes from unintentionally looking only at a certain proportion of things or people that have been through some process (Shermer 2014). We hear much more from the '*winners*' (celebrities, billionaires, Olympic champions, etc.) than the '*non-winners*' (those who are not celebrities or billionaires or Olympic champions, etc.) about how to win" (emphasis added). Accordingly, one might worry that the 32 women philosophers listed in Table 3 are "winners," i.e., the most cited women philosophers in the SEP, who have fought in the "argumentative arena" (Alcoff 2013) of academic philosophy and survived, but there are plenty of other women who have fought in the "argumentative arena" (Alcoff 2013) of academic philosophy as well but did not survive. Those women are not represented in our sample of women philosophers.

We acknowledge this point.[8] However, we still think that it would be interesting to find out if there are any significant differences between the types of arguments advanced in the published works of the most cited male philosophers and the types of arguments advanced in the published works of the most cited women philosophers, even if our sample of women philosophers includes "elite" women philosophers who have managed to survive the "argumentative arena" (Alcoff 2013) of academic philosophy. For if we were to find significant differences in patterns of argumentation between the most cited male philosophers and the most cited women philosophers, then such findings could provide at least prima facie empirical

[8] Many thanks to an anonymous reviewer for raising this point.

support to “the supposition that, philosophy being concerned above all with arguments, women are naturally less adept in the field” (Warnock 2015). On the other hand, if we were to find no significant differences in patterns of argumentation between the most cited male philosophers and the most cited women philosophers, then such findings could suggest that women philosophers *can* be just as concerned with arguments, and just as philosophically argumentative, as male philosophers supposedly are.

2.2 *Text Mining Methods*

A combination of several text-mining packages in R Language were used to manipulate the corpus of philosophical texts (n = 4,408) throughout this study. The publication years for documents in this corpus ranged from 1874 to 2017. RStudio was used as an interactive-development environment to process the data. The corpus of documents included a .txt file containing the full-text of the philosophical works, and a corresponding .xml file to the full-text file composed of the metadata information about each text file.

The *readtext* package was utilized to load the text files into the RStudio environment. The *readtext* function takes a folder path as an input parameter (i.e., readtext(“filepath”). The *readtext()* function will then load all files in the target folder into RStudio as a dataframe. The dataframe will consist of two columns. The first column is titled “doc_id” and it lists the file names as individual elements within a string vector. The second column is titled “text” and it includes the full-text from each of the individual text files as a single character string. The result is a vector of character strings, with each string containing the full-text of an input text file. The .xml files were converted to .txt files from the Windows Command Prompt application and also read into R using the *readtext()* function.

Windows Command Prompt to change file type to .txt:

```
>>>  cd (folderpath)
>>> ren *.txt *.xml
```

R language example of *readtext()*:

```
# The full_text and metadata objects will load the .txt files as a
data frame consisting of the document ID and the full_text.
> full_text = readtext("C:/Users/mdick/Desktop/full_text_documents")
> metadata = readtext("C:/Users/mdick/Deskteop/metadata")

> glimpse(full_text)
Rows: 435,703
Columns: 2
$ doc_id <chr> "book-chapter-10.1163_j.ctt1w8h0ph.11.txt"~
$ text   <chr> "<plain_text> <page sequence=\"1\"> CHAPTER 8 On
Statehood  Failed Stat~

> glimpse(metadata)
Rows: 466,486
Columns: 2
$ doc_id <chr> "book-chapter-10.1163_j.ctt1w8h0ph.11.txt", "book-
chapter-10.1163_j.ctt~
$ text   <chr> "<book xmlns:oasis=\"http://docs.oasis-open.org/ns/~
```

To search for indicator pairs within the full-text documents, the *string_detect()* function from the *stringr* package was used in combination with a regular expression (regex) as a pattern search parameter. The argument indicator root and anchor were included within the regular expression to search for specific words. The general pattern of the regular expression follows:

```
# A regular expression to detect the root and anchor words within the
specified word-range
> pattern_matches = corpus[ 0, 2] %>% str_detect(
regex("(?:proves\\W+(?:\\w+\\W+){0,3}?probably|probably\\W+(?:\\w+\\W
```

```
+){0,3}?proves)"))

# View the beginning of the logicals list
> head(pattern_matches)
[1] FALSE FALSE FALSE FALSE FALSE FALSE
```

In the regular expression, the "{0,3}" represents the 3-word range limit, exclusively. The 3 was replaced with 6 or 10 when searching across those respective word-ranges. In the regex, "root" and "anchor" represent the respective halves of the indicator pair. These pairs were changed and pattern search was run across the entire corpus to generate a list of the articles which contained matches. While it may not seem intuitive to search for root-anchor argument indicator pairs across the whole corpus, this technique was employed in an earlier study (Mizrahi and Dickinson 2021). It was a simpler task to filter through all articles containing matches, rather than to drill down to the authors used in this study and then compile the matched articles.

The regular expression pattern allows for the root of the argument indicator pairs to both precede and follow the anchor word(s) within a certain range of words, exclusively. The function was applied to the corpus across three word-ranges. The ranges selected permitted 3, 6, or 10 words between the argument indicator root and the anchor word(s). For example, to search for pattern matches across a range of 3 words, the regular expression returns a positive match in the following cases:

Root $word_1$ $word_2$ $word_3$ *Anchor* | OR | *Anchor* $word_1$ $word_2$ $word_3$ *Root*

Any pattern in which the argument indicator roots and anchors are separated by less than the maximum range (i.e., 3, 6, or 10) is also considered a positive match. For example, as applied within a 3-word maximum range, the following case would be considered a positive match:

Anchor $word_1$ $word_2$ *Root*

Applied in this manner, the *string_detect()* function will return a list of TRUE or FALSE logical values, where TRUE indicates the presence of the argument indicator and the anchor at least one time within each document and FALSE indicates no pattern match. The logical values were then converted to numeric data, with 1 replacing TRUE and 0 replacing FALSE.

```
# View the converted list of numbers
> head(log_to_num)
[1] 0 0 0 0 0 0

# Bind the list of numbers to the corpus
> corpus_w_matches = cbind(corpus, log_to_num)

# View the new data frame with the numeric list included
> glimpse(corpus_w_matches)
Rows: 435,703
Columns: 3
$ doc_id     <chr> "book-chapter-10.1163_j.ctt1w8h0ph.11.txt"~
$ text       <chr> "<plain_text> <page sequence=\"1\"> CHAPTER 8 On~
$ log_to_num <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~

# Filter for matched articles (i.e. log_to_num = 1)
> pos_match_iso = corpus_w_matches %>%
+   filter( == 1)

# Remove the corpus_w_matches
> pos_match_iso$log_to_num = NULL

# View the positively matched articles containing the indicator
# pairing "proves probably" within 3 words, exclusively.
> glimpse(pos_match_iso)
Rows: 5
Columns: 2
$ doc_id <chr> "journal-article-10.2307_20128119.txt"_
$ text   <chr> "<plain_text><page sequence=\"1\">WHY FIRE GOES UP:~
```

The numbers list is then bound to the corpus using the *cbind()* function. The new corpus is then filtered for the positively matched articles. This detection process was repeated for each indicator pair of the deductive, inductive, and abductive lists (see Table 2) and across all three word-ranges. The resulting lists were then summed, and the number of positive matches were recorded to a separate .csv file.

Separate .csv files containing matched full-text documents compiled by word-range were also generated from these lists. Lists were compiled by argument type and labeled in a new column. From these newly generated lists of matched articles, a series of regular expressions was used to isolate XML tags containing the author names for each document from the metadata files. The extracted names were then joined with the corresponding positively matched full-text files. Additional metadata was pulled, such as publication year, and journal title. The *filter()* command from the dplyr package is case-sensitive, so authors' given names and surnames were lowered to allow for easier filtering.

```
# Extract the surnames for the authors from the metadata
> surname_w_html = corpus$metadata %>%
  str_extract("surname>(.*?)<")

# View the messy extracted data
> head(surname_w_html)
[1] "surname>Spaulding<" "surname>Kaufmann<"  "surname>Aaron<"
[4] "surname>Clarke<"    "surname>JACKSON<"   "surname>Carus<"


# Remove the HTML tags from the extracted text
> surname_no_html = surname_w_html %>%
  str_remove("surname>") %>%
  str_remove("<")


# View the cleaned extracted text
```

```
> head(surname_no_html)
[1] "Spaulding" "Kaufmann"  "Aaron"     "Clarke"    "JACKSON"
[6] "Carus"
```

The matched articles were then filtered among the 64 most widely cited authors' works from the 32 men and 32 women philosophers examined in this study (see Table 3). It is important to note that filtering on the authors' given names posed a unique challenge, as some authors have multiple listings for their first names. Martha Nussbaum, for example, can be listed as "martha", "martha c.", "m. c.", and "martha craven" as a given name. While authors' given names would have needed to be filtered individually anyway, the presence of multiple labels for the given_name column required close attention to detail to ensure that the wrong authors were not being taken into consideration.

```
# Load the compiled list of articles containing matched indicator
pairs and extracted metadata
> data_3 = read.csv("meta_extracted_master_3.csv")

# View a snapshot of the data with the metadata extracted
Rows: 49,315
Columns: 10
$ doc_id        <fct> journal-article-10.2307_2011251~
$ text.x        <fct> "<plain_text><page sequence=\"1~
$ arg_type      <fct> deductive, deductive, deductive~
$ text.y        <fct> "<article xmlns:xsi=\"http://ww~
$ title         <fct> "The Journal of Philosophy, Psy~
$ subtitle      <fct> NA, NA, NA, NA, NA, NA, NA, NA,~
$ journal_title <fct> "The Journal of Philosophy, Psy~
$ surname       <chr> "spaulding", "kaufmann", "aaron~
$ given_name    <chr> "edward g.", "felix", "r. i.", ~
$ year          <int> 1906, 1944, 1942, 1959, 1990, 1~

# Filter by authors' surname to see all authors with the specific
```

```
# surname
> nussbaum = data_3 %>% filter(surname == "nussbaum")

# View the first names of the authors with the specified surname
> nussbaum$given_name
 [1] "martha c." "martha c." "martha c." "martha c." "martha c."
[6] "martha c." "charles" "martha" "martha c." "martha c."
[11] "martha c." "martha c." "martha c." "martha c." "martha c."
[16] "martha c." "martha c." "martha" "martha c." "martha c."
[21] "martha c." "martha c." "martha" "martha c." "martha c."
[26] "martha c." "martha c." "martha c." "martha c." "martha"
[31] "martha c." "charles"
# Filter by the specified last name and authors' various first names
to remove unwanted authors who share the same surname
> m_nussbaum = data_3 %>% filter(surname == "nussbaum",
                                 given_name == "martha" |
                                 given_name == "martha c." |
                                 given_name == "m. c." |
                                 given_name == "martha craven")
> m_nussbaum$given_name
 [1] "martha c." "martha c." "martha c." "martha c." "martha c."
[6] "martha c." "martha" "martha c." "martha c." "martha c."
[11] "martha c." "martha c." "martha c." "martha c." "martha c."
[16] "martha c." "martha"    "martha c." "martha c." "martha c."
[21] "martha c." "martha" "martha c." "martha c." "martha c."
[26] "martha c." "martha c." "martha c." "martha" "martha c."
```

The full-text string vectors were then stripped of any HTML tags from the full-text strings using more regex. This was done to produce a more accurate word count. The *wordcount()* function will ignore case but will designate a space in the text as the break between words. This means that the HTML tags would have been counted and resulted in less accurate totals for each author. One issue with stripping the HTML and other formatting centered on the use of the forward slash "/" to indicate a new line in the text. These standalone forward slashes would have been counted originally and proved difficult to remove from the text. Ultimately, the

space between each standalone forward slash and the previous word was removed, appending the forward slash to the end of the previous word and eliminating each floating forward slash from the overall word count. For example, a string vector of “word1, word2, word3/” would return a word count of 3, ignoring the numbers and the forward slash, as they are not separated by a blank space.

```
View a snapshot of the text in m_nussbaum
> glimpse(m_nussbaum$text.x)
 chr [1:30] "<plain_text><page sequence=\"1\">Â© Metaphilosophy LLC
and Blackwell Publishers Ltd. 2002. Published by Blackwe"|
__truncated__ ...

# A function to strip the HTML tags from the full-text strings
> strip_html = function(data) {

        data$text.x %>%
                str_remove_all("<[^>]*>") %>%
                str_remove_all("\n")
}

# Apply the strip_all() function to the filtered Martha Nussbaum data
> m_nussbaum_no_html = strip_all(m_nussbaum)

# View the text after the HTML tags have been stripped
> glimpse(m_nussbaum_no_html)
 chr [1:30] "Â© Metaphilosophy LLC and Blackwell Publishers Ltd.
2002. Published by Blackwell Publishers, 108 Cowley Road, O"|
__truncated__ ...

# Count the number of characters in the filtered full-text
> wordcount(m_nussbaum_no_html)
[1] 435349
```

Due to how the matched articles were compiled, the datasets contained duplicate records. These duplicates were removed using the distinct() function to ensure each matched record would only be listed one time in the dataset and the total word count for each author from across the entire corpus was then calculated for each of the 64 authors. The articles were then filtered by matched argument-type (i.e., deductive, inductive, and abductive). Word counts were then calculated for each author across each argument type and compiled into a master .csv file of author word counts.

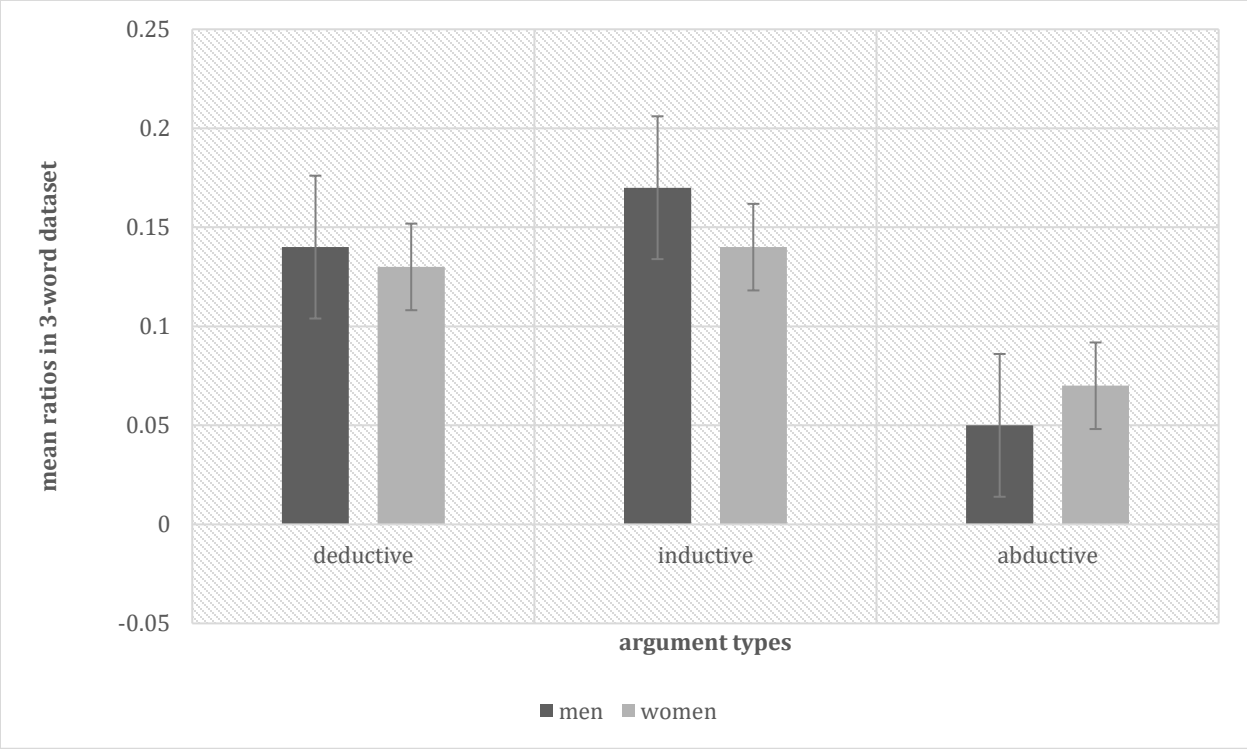
It should be noted that the algorithm searches the entire corpus for each indicator pair but can only match each indicator pair with a single article one time. For example, if 'therefore necessarily' and 'hence certainly' occur in article *x* published in 1950, the algorithm would return a count of 2. This would be the case even if, say, 'hence certainly' repeated more than once throughout article *x*. If article *y* was also published in 1950 and contains the same indicator pairs, they would also be counted, and the algorithm would return a count of 4 for 1950.

It is important to emphasize again that this search algorithm is not totally immune from counting false negatives and/or false positives, as we discussed in Section 2.1. We did, however, test the algorithm on small amounts of textual data. Once the algorithm could count the correct number of exact matches to the provided root-anchor pairings on small amounts of data, it was scaled up to run on larger sections of the corpus, and eventually on the entire corpus at once. Another limitation is that processing times for each indicator pair could range from 10-30 minutes within the RStudio application. RStudio could also potentially run out of RAM while loading in digital objects. However, repeated tests did result in the same number of matches per indicator-pair and across each word-range.

3. Results

Since some academic philosophers publish more than others, we need to calculate the ratios or proportions of each argument type (see Table 2) in the published works of men and women philosophers (see Table 3). Then we can compare the mean ratios in order to find out whether there are any statistically significant differences between the types of arguments made by men and women philosophers in academic philosophy. In searches permitting three words between argument indicator root and anchor, we have found no statistically significant differences between the types of arguments made by men and women philosophers. See Figure 1. (For a complete list of ratios by philosopher in our 3-word dataset, see Appendix I.)

Figure 1. Mean ratios of argument types made by men and women philosophers (3-word dataset)



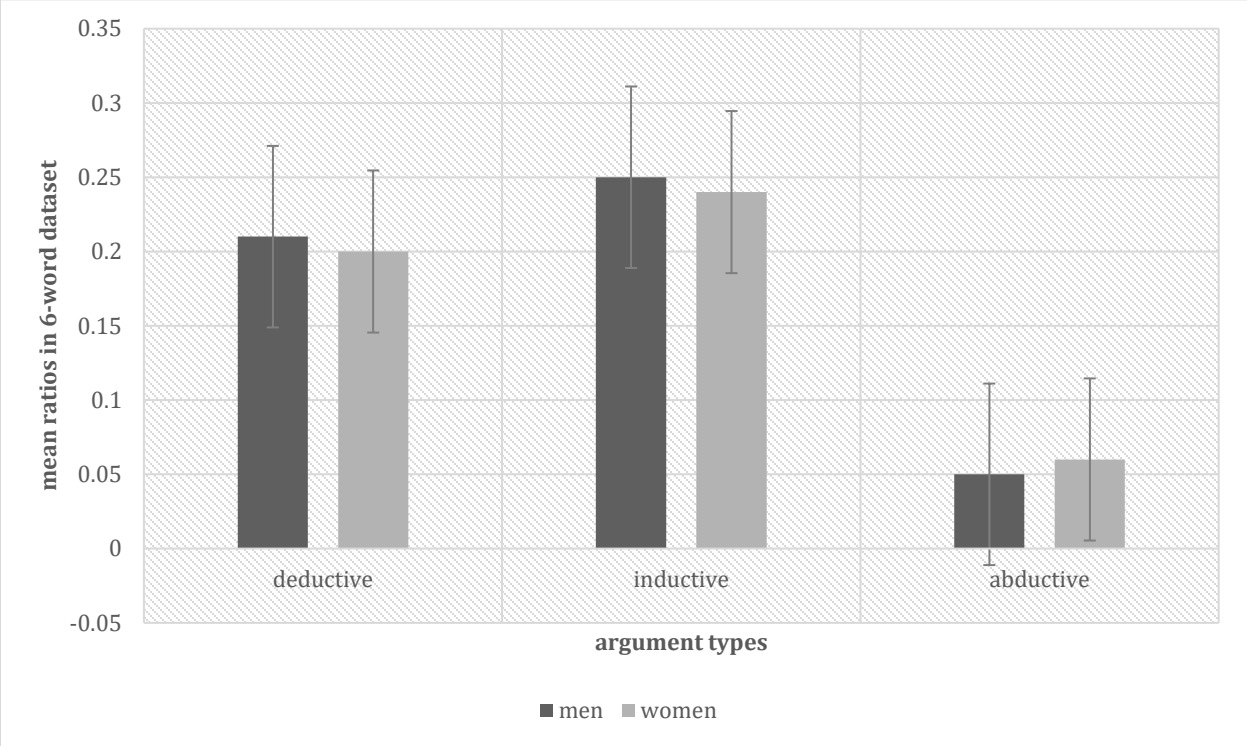
An independent-samples t-test was conducted to compare the ratios of deductive arguments in works written by male philosophers and works written by women philosophers

from the results for searches allowing a 3-word maximum range. There was no significant difference between the ratios of deductive arguments in works written by male philosophers (M = 0.14, SD = 0.08, N = 32) and the ratios of deductive arguments in works written by women philosophers (M = 0.13, SD = 0.11, N = 32), $t(59) = 0.25$, $p = 0.79$, two-tailed. Likewise, an independent-samples t-test was conducted to compare the ratios of inductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 3-word maximum range. There was no significant difference between the ratios of inductive arguments in works written by male philosophers (M = 0.17, SD = 0.1, N = 32) and the ratios of inductive arguments in works written by women philosophers (M = 0.14, SD = 0.14, N = 32), $t(57) = 1.12$, $p = 0.26$, two-tailed. Finally, an independent-samples t-test was conducted to compare the ratios of abductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 3-word maximum range. There was no significant difference between the ratios of abductive arguments in works written by male philosophers (M = 0.05, SD = 0.05, N = 32) and the ratios of abductive arguments in works written by women philosophers (M = 0.07, SD = 0.1, N = 32), $t(49) = -0.81$, $p = 0.41$, two-tailed. These results suggest that women philosophers make deductive, inductive, and abductive arguments in their published works just as much as male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work.

As we did with our 3-word dataset, we calculated the ratios or proportions of each argument type (see Table 2) in the published works of men and women philosophers (see Table 3) using data from our 6-word dataset. In searches permitting six words between argument indicator root and anchor, as in searches permitting three words between argument indicator root

and anchor, we have found no statistically significant differences between the types of arguments made by men and women philosophers. See Figure 2. (For a complete list of ratios by philosopher in our 6-word dataset, see Appendix II.)

Figure 2. Mean ratios of argument types made by men and women philosophers (6-word dataset)

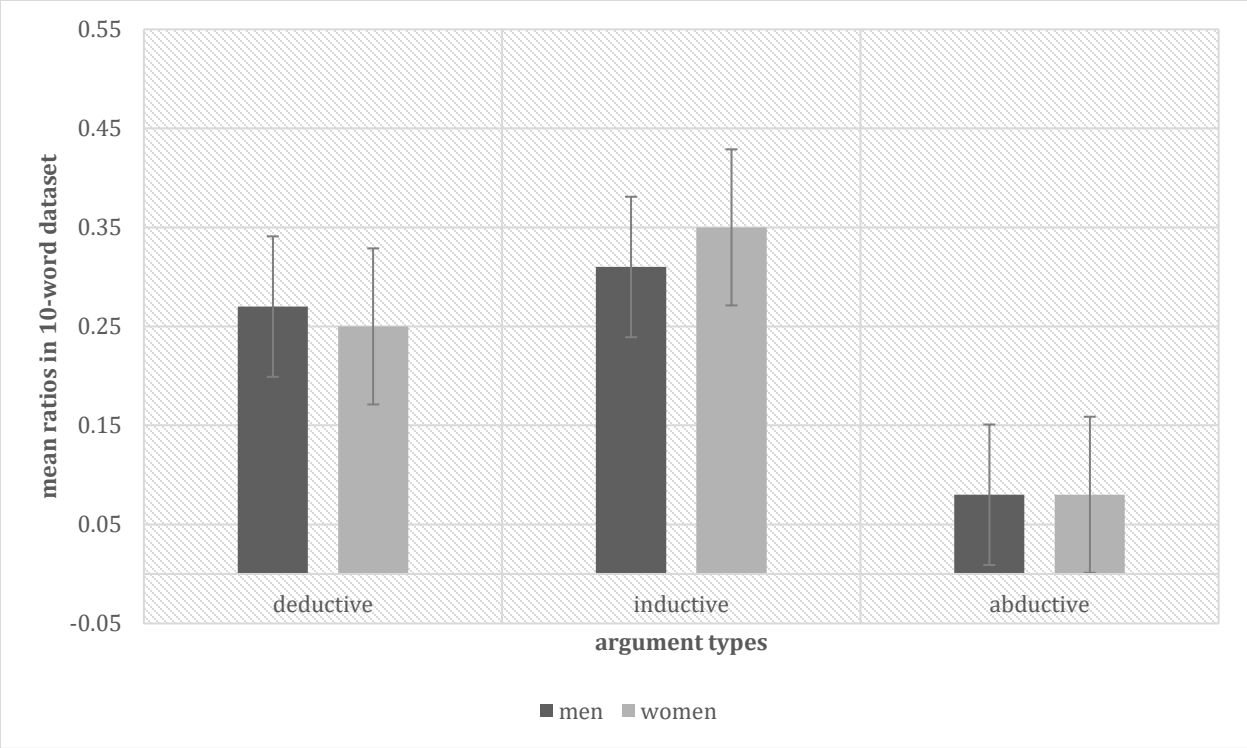


An independent-samples t-test was conducted to compare the ratios of deductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 6-word maximum range. There was no significant difference between the ratios of deductive arguments in works written by male philosophers (M = 0.21, SD = 0.12, N = 32) and the ratios of deductive arguments in works written by women philosophers (M = 0.2, SD = 0.15, N = 32), $t(60) = 0.39$, $p = 0.69$, two-tailed. Likewise, an independent-samples t-test was conducted to compare the ratios of inductive arguments in works written by male philosophers and works written by women philosophers from the results for

searches allowing a 6-word maximum range. There was no significant difference between the ratios of inductive arguments in works written by male philosophers (M = 0.25, SD = 0.13, N = 32) and the ratios of inductive arguments in works written by women philosophers (M = 0.24, SD = 0.18, N = 32), $t(56) = 0.16$, $p = 0.86$, two-tailed. Finally, an independent-samples t-test was conducted to compare the ratios of abductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 6-word maximum range. There was no significant difference between the ratios of abductive arguments in works written by male philosophers (M = 0.05, SD = 0.05, N = 32) and the ratios of abductive arguments in works written by women philosophers (M = 0.06, SD = 0.07, N = 32), $t(56) = -0.65$, $p = 0.51$, two-tailed. These results, which are consistent with the results obtained from our 3-word searches, suggest that women philosophers make deductive, inductive, and abductive arguments in their published works just as much as male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work.

As we did with our 3-word and 6-word datasets, we calculated the ratios or proportions of each argument type (see Table 2) in the published works of men and women philosophers (see Table 3) using data from our 10-word dataset. In searches permitting ten words between argument indicator root and anchor, as in searches permitting three words and those permitting six words between argument indicator root and anchor, we have found no statistically significant differences between the types of arguments made by men and women philosophers. See Figure 3. (For the complete list of ratios by philosopher in our 10-word dataset, see Appendix III.)

Figure 3. Mean ratios of argument types made by men and women philosophers (10-word dataset)



An independent-samples t-test was conducted to compare the ratios of deductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 10-word maximum range. There was no significant difference between the ratios of deductive arguments in works written by male philosophers (M = 0.27, SD = 0.11, N = 32) and the ratios of deductive arguments in works written by women philosophers (M = 0.27, SD = 0.16, N = 32), $t(57) = 0.47$, $p = 0.63$, two-tailed. Likewise, an independent-samples t-test was conducted to compare the ratios of inductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 10-word maximum range. There was no significant difference between the ratios of inductive arguments in works written by male philosophers (M = 0.31, SD = 0.15, N = 32) and the ratios of inductive arguments in works written by women philosophers (M = 0.35, SD = 0.24, N = 32), $t(52) = -0.75$, $p = 0.45$, two-tailed. Finally, an independent-samples t-test

was conducted to compare the ratios of abductive arguments in works written by male philosophers and works written by women philosophers from the results for searches allowing a 10-word maximum range. There was no significant difference between the ratios of abductive arguments in works written by male philosophers (M = 0.08, SD = 0.07, N = 32) and the ratios of abductive arguments in works written by women philosophers (M = 0.08, SD = 0.09, N = 32), $t(58) = -0.21$, $p = 0.82$, two-tailed. These results, which are consistent with the results obtained from our 3-word and 6-word searches, suggest that women philosophers make deductive, inductive, and abductive arguments in their published works just as much as male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work.

4. Discussion

As we discussed in Section 1 above, some philosophers have suggested that it may be the methods, specifically, the "logic-chopping" and "paradox-mongering" concern with arguments (Friedman 2013, p. 25), of academic philosophy that explain the underrepresentation of women in the discipline (Demarest et al. 2017, p. 530). By contrast, Warnock (2015) claims that academic philosophy's "women problem" cannot "be explained by the supposition that, philosophy being concerned above all with arguments, women are naturally less adept in the field."

Accordingly, our quantitative, corpus-based study was designed to address the following research questions empirically:

1. Do men and women philosophers make different types of arguments in their published works?

2. If men and women philosophers make different types of arguments in their published works, which types of arguments are typically made by male philosophers and which types of arguments are typically made by women philosophers? Are there significant differences between the types of arguments typically made by male philosophers versus those typically made by women philosophers?

The results of our empirical study suggest the following tentative answers to these research questions. Our results suggest that both men and women philosophers make arguments in their published works. More specifically, our data reveal no statistically significant differences between the types of arguments advanced in published works written by male philosophers and the types of arguments advanced in published works written by women philosophers. In fact, both men and women philosophers make the three types of arguments we have searched for systematically, namely, deductive arguments, inductive arguments, and abductive arguments, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work. Since we have observed these patterns in our 3-word, 6-word, and 10-word datasets, we can be quite confident that these results are robust.

Now, it would be premature to draw any sweeping conclusions from the results of our empirical study vis-a-vis the "gender problem" or "women problem" in academic philosophy because of the so-called "survivor bias" in our sample of women philosophers. That is, as we discussed in Section 2.1 above, the 32 women philosophers in our sample of women philosophers could be considered "elite" women philosophers who have managed to survive the "argumentative arena" (Alcoff 2013) of academic philosophy. However, as we also discussed in Section 2.1 above, we think it is still interesting to find out that there are no statistically significant differences in patterns of argumentation found in the published works of the most

cited male philosophers and those found in the published works of the most cited women philosophers. For the absence of statistically significant differences in patterns of argumentation between the most cited male philosophers and the most cited women philosophers suggests that women philosophers *can* be just as concerned with arguments, and just as philosophically argumentative, as male philosophers supposedly are.

Even though we did not find statistically significant differences in patterns of argumentation between the most cited male philosophers and the most cited women philosophers, these empirical findings should not be construed as conclusive evidence against "the supposition that, philosophy being concerned above all with arguments, women are naturally less adept in the field" (Warnock 2015). For, as we discussed in Section 2.1 above, argument indicators are reliable indicators, but they are not "sure signs of whether an argument is present" (Lepore and Cumming 2013, p. 6), and so our search algorithm is not totally immune from counting false negatives and/or false positives. Moreover, as we discussed in Section 1 above, while academic publications are one aspect of the "argumentative arena" (Alcoff 2013) of academic philosophy, there are other aspects as well. More explicitly, while professional philosophers do make arguments in academic publications, they also make arguments in academic conferences, academic seminars, and the like. Therefore, it could be argued that, while it may be interesting to have empirical evidence suggesting that the most cited women philosophers make deductive, inductive, and abductive arguments in their published works just as much as the most cited male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work, such empirical evidence gives us only a partial picture of the "argumentative arena" (Alcoff 2013) in academic philosophy. Indeed, it could be argued that "the argument-as-war metaphor," which is meant to

describe academic philosophy's culture of aggressive argument (Rooney 2010), refers to "face-to-face arguments" in particular rather than to arguments in general (Beebee 2013, p. 67).

As we discussed in Section 1 above, our empirical study is focused on arguments made in academic publications specifically, rather than face-to-face arguments. Clearly, then, our results cannot be extended to those other aspects of academic philosophy's "argumentative arena" (Alcoff 2013), such as the academic conference or the philosophy seminar (Beebee 2013, pp. 63-73). In addition to seminars and conferences, it might also be interesting to examine rejected papers as well. For it may be the case that papers written by women philosophers were rejected by philosophy journals precisely because they were deemed "insufficiently argumentative." Furthermore, as an anonymous reviewer suggested, future work with data gathered from published books might also reveal patterns of argumentation that are different from those we have found in published articles. For these reasons, we submit, further studies are needed to better understand these other aspects of academic philosophy's "argumentative arena" (Alcoff 2013), beyond arguments made in academic journals, so as to be able to ascertain whether they play a role in explaining the gender gap in academic philosophy.

5. Conclusion

Some philosophers have suggested that it may be the methods, specifically, the "logic-chopping" and "paradox-mongering" concern with arguments (Friedman 2013, p. 25), of academic philosophy that explain the underrepresentation of women in the discipline (Demarest et al. 2017, p. 530). By contrast, Warnock (2015) claims that academic philosophy's "women problem" cannot "be explained by the supposition that, philosophy being concerned above all with arguments, women are naturally less adept in the field." Our quantitative, corpus-based

study was designed to contribute to the literature on the gender gap in academic philosophy by detecting patterns of argumentation in the published works of 32 men and 32 women of the most cited academic philosophers in the SEP. Overall, the results of our empirical study suggest that women philosophers make deductive, inductive, and abductive arguments in their published works just as much as male philosophers do, with no statistically significant differences in the proportions of those arguments relative to each philosopher's body of work. These results suggest that women philosophers can be just as concerned with arguments, and just as philosophically argumentative, as male philosophers supposedly are. Since arguments made in academic publications are but one aspect of academic philosophy's "argumentative arena" (Alcoff 2013), which include face-to-face arguments made in philosophy conferences and seminars as well, we think that further studies are needed to investigate these other aspects of the "argumentative arena" (Alcoff 2013) in academic philosophy in order to ascertain whether they play a role in explaining the gender gap in academic philosophy.

Acknowledgements

We are very grateful to two anonymous reviewers of *Synthese* for their helpful comments on earlier drafts of this paper.

Declarations

Funding: not applicable

Conflict of interest: the authors declare no conflicts of interest

Availability of data: data are available

Code availability: code is available

References

Adleberg, Toni, Morgan Thompson, and Eddy Nahmias. (2015). Do Men and Women have Different Philosophical Intuitions? Further Data. *Philosophical Psychology* 28 (5): 615-641.

Alcoff, Linda. (2013). What's wrong with philosophy? *The New York Times*, September 3, 2013. https://opinionator.blogs.nytimes.com/2013/09/03/whats-wrong-with-philosophy/.

Antony, Louise. (2012). Different Voices or Perfect Storm: Why are There So Few Women in Philosophy? *Journal of Social Philosophy* 43 (3): 227-255.

Ashton, Zoe and Moti Mizrahi. (2018). Show Me that Argument: Empirically Testing the Armchair Philosophy Picture. *Metaphilosophy* 49 (1-2): 58-70.

Baronett, Stan. (2016). *Logic*. New York: Oxford University Press.

Beasley, Eric. (2018). *Misperceptions of the Social World: What We Get Wrong about Sex, Race, Money, and Violence*. New York: Routledge.

Beebee, Helen. (2013). Women and Deviance in Philosophy. In K. Hutchison and F. Jenkins (eds.), *Women in Philosophy: What Needs to Change?* (pp. 61-80). New York: Oxford University Press.

Bowell, Tracy. (2015). The Problem(s) of Women in Philosophy: Reflections on the Practice of Feminism in Philosophy from Contemporary Aotearoa/New Zealand. *Women's Studies Journal* 29 (2): 4-21.

Buckwalter, Wesley and Stephen Stich. (2014). Gender and Philosophical Intuition. In Joshua Knobe and Sean Nichols (eds.), *Experimental Philosophy* (Vol. 2, 307–346). Oxford: Oxford University Press.

Currie, Gregory. (2016). Does Great Literature Make Us Better? In P. Catapano and S. Critchley (eds.), *The Stone Reader: Modern Philosophy in 133 Arguments* (pp. 198-202). New York: W. W. Norton & Co.

Demarest, Heather, Seth Robertson, Megan Haggard, Madeline Martin-Seaver, and Jewelle Bickel. (2017). Similarity and Enjoyment: Predicting Continuation for Women in Philosophy. *Analysis* 77 (3): 525-541.

Easton, Christina. (2021). Women and 'the philosophical personality': evaluating whether gender differences in the Cognitive Reflection Test have significance for explaining the gender gap in Philosophy. *Synthese* 198 (1): 139-167.

Friedman, Marylin. (2013). Women in Philosophy: Why Should We Care? In Katrina Hutchison and Fiona Jenkins (eds.), *Women in Philosophy: What Needs to Change?* (pp. 21-38). New York: Oxford University Press.

Govier, Trudy. (2013). *A Practical Study of Argument*. Enhanced Seventh Edition. Belmont, CA: Wadsworth, Cengage Learning.

Harrell, Maralee. (2016). *What is the Argument? An Introduction to Philosophical Argument and Analysis*. Cambridge, MA: The MIT Press.

Hurley, P. J. and Watson, L. (2018). *A Concise Introduction to Logic*. Thirteenth Edition. Boston, MA: Cengage Learning.

Krishnamurthy, Meena, Shen-yi Liao, Monique Deveaux, and Maggie Dalecki. (2017). The Underrepresentation of Women in Prestigious Ethics Journals. *Hypatia* 32 (4): 928-939.

Lepore, Ernest and Sam Cumming. 2013. *Meaning and Argument: An Introduction to Logic through Language*. Malden, MA: Blackwell.

Lewis, Brooke. (2009). Where Are All the Women? *The Philosophers' Magazine* 47 (4): 12-16.

Maki, Uskali. (2009). MISSing the World. Models as Isolations and Credible Surrogate Systems. *Erkenntnis* 70 (1): 29-43.

Marcus, Russell. (2018). *Introduction to Formal Logic with Philosophical Applications*. New York: Oxford University Press.

McGinn, Colin. (2007). Review of Ted Honderich's *On Consciousness*. *The Philosophical Review* 116 (3): 474-477.

Mizrahi, M. and Mike Dickinson. (2021). The Analytic-Continental Divide in Philosophical Practice: An Empirical Study. *Metaphilosophy* 52 (5): 668-680.

Morrow, D. R. and Anthony Weston. (2019). *A Workbook for Arguments: A Complete Course in Critical Thinking*. Third Edition. Indianapolis: Hackett Publishing Co.

Overton, J. A. (2013). "Explain" in Scientific Discourse. *Synthese* 190 (8): 1383-1405.

Paxton, Molly, Carrie Figdor, and Valerie Tiberius. (2012). Quantifying the Gender Gap: An Empirical Study of the Underrepresentation of Women in Philosophy. *Hypatia* 27 (4): 949-957.

Rooney, Phyllis. (2010). Philosophy, Adversarial Argumentation, and Embattled Reason. *Informal Logic* 30 (3): 203-234.

Schwitzgebel, Eric. (2019). The 295 Most-Cited Contemporary Authors in the Stanford Encyclopedia of Philosophy. *The Splintered Mind*, August 20, 2019. http://schwitzsplinters.blogspot.com/2019/08/the-295-most-cited-contemporary-authors.html.

Schwitzgebel, Eric and C. D. Jennings. (2017). Women in Philosophy: Quantitative Analyses of Specialization, Prevalence, Visibility, and Generational Change. *Public Affairs Quarterly* 31: 83-105.

Shermer, Michael. (2014). Surviving Statistics. *Scientific American* 311 (3): 94.

Sober, Elliot. (2015). Two Cornell Realisms: Moral and Scientific. *Philosophical Studies* 172 (4): 905-924.

Thompson, Morgan, Toni Adleberg, Sam Sims, and Eddy Nahmias. (2016). Why Do Women Leave Philosophy? Surveying Students at the Introductory Level. *Philosophers' Imprint* 16 (6): 1-36.

Thompson, Morgan. (2017). Explanations of the gender gap in philosophy. *Philosophy Compass* 12: e12406.

Toulmin, S. E. (1976). *Knowing and Acting: An Invitation to Philosophy*. New York: Macmillan.

Walton, Douglas. (1999). *One-Sided Arguments: A Dialectical Analysis of Bias*. Albany: State University of New York Press.

Warnock, Mary and Julian Baggini. (2015). Does philosophy have a problem with women? *The Guardian*, July 25, 2015. https://www.theguardian.com/commentisfree/2015/jul/25/philosphy-women-warnock-baggini-debate.

Wilhelm, Isaac, Sherri Lynn Conklin, and Nicole Hassoun. (2018). New data on the representation of women in philosophy journals: 2004–2015. *Philosophical Studies* 175 (6): 1441-1464.

Wolf, Susan. (1986). Above and Below the Line of Duty. *Philosophical Topics* 14 (2): 131-148.

Wuest, Amy. (2013). Yes, There Is a Problem: What Is to Be Done about the Climate for Women in Philosophy? *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 3 (1): 146-150.

Appendix I. Counts and ratios of argument types by philosopher in the 3-word dataset

Philosopher	**Total**	**De count**	**De ratio**	**In count**	**In ratio**	**Ab count**	**Ab ratio**
Lewis, David K.	295584	12753	0.04	30621	0.10	0	0.00
Quine, W.V.O.	453628	69650	0.15	19941	0.04	0	0.00
Putnam, Hilary	498836	64165	0.13	102324	0.21	35917	0.07

Rawls, John	193515	11957	0.06	75723	0.39	0	0.00
Davidson, Donald	238097	40014	0.17	17964	0.08	9406	0.04
Kripke, Saul	70655	7124	0.10	29528	0.42	0	0.00
Williams, Bernard	229423	76420	0.33	39974	0.17	17456	0.08
Nozick, Robert	84856	2412	0.03	22966	0.27	0	0.00
Williamson, Timothy	574871	125147	0.22	41636	0.07	15724	0.03
Jackson, Frank	456979	51701	0.11	55961	0.12	24098	0.05
Nagel, Thomas	172782	13858	0.08	38077	0.22	14180	0.08
Searle, John R.	247454	11335	0.05	30105	0.12	17863	0.07
Van Fraassen, Bas	492842	89874	0.18	56295	0.11	13415	0.03
Armstrong, David M.	274647	46657	0.17	48040	0.17	9585	0.03
Dummett, Michael	295445	100750	0.34	62752	0.21	18666	0.06
Fodor, Jerry	317155	66482	0.21	83310	0.26	10257	0.03
Harman, Gilbert	348876	30815	0.09	59148	0.17	65333	0.19
Chisholm, Roderick	479455	34884	0.07	22395	0.05	0	0.00
Dennett, Daniel C.	294155	44146	0.15	55117	0.19	50282	0.17
Chalmers, David J.	213641	56619	0.27	73261	0.34	49959	0.23
Strawson, P.F.	341154	56197	0.16	25103	0.07	21086	0.06

Stalnaker, Robert	389202	13408	0.03	18872	0.05	20989	0.05
Scanlon, T.M.	182519	19072	0.10	20119	0.11	4239	0.02
Dworkin, Ronald	154721	25716	0.17	44182	0.29	9085	0.06
Pettit, Philip	465236	55621	0.12	149256	0.32	0	0.00
Fine, Kit	620917	188554	0.30	28776	0.05	34012	0.05
Sober, Elliott	528653	44523	0.08	67028	0.13	20782	0.04
Van Inwagen, Peter	455873	97816	0.21	94439	0.21	28458	0.06
Popper, Karl	270095	44931	0.17	32727	0.12	0	0.00
Parfit, Derek	181861	0	0.00	52928	0.29	0	0.00
Kitcher, Philip	538719	32999	0.06	58543	0.11	76636	0.14
Bennett, Jonathan	415384	66990	0.16	72266	0.17	11818	0.03
Nussbaum, Martha	476306	77325	0.16	221079	0.46	0	0.00
Anscombe, G.E.M.	129188	17503	0.14	0	0.00	0	0.00
Korsgaard, Christine	152127	41033	0.27	15783	0.10	0	0.00
Anderson, Elizabeth	216245	26240	0.12	27160	0.13	10830	0.05
Thomson, Judith Jarvis	305937	67227	0.22	59026	0.19	0	0.00
Cartwright, Nancy	256543	4962	0.02	23823	0.09	12924	0.05
Annas, Julia	245770	48301	0.20	32110	0.13	15863	0.06

Young, Iris Marion	141579	19189	0.14	31872	0.23	19934	0.14
Millikan, Ruth G.	250734	8535	0.03	72808	0.29	48753	0.19
Foot, Philippa	122136	49478	0.41	0	0.00	0	0.00
Stump, Eleonore	200811	53264	0.27	12416	0.06	31434	0.16
Okin, Susan Moller	160664	33343	0.21	85629	0.53	0	0.00
Butler, Judith	22587	0	0.00	0	0.00	0	0.00
O'Neill, Onora	134399	0	0.00	5792	0.04	0	0.00
Zagzebski, Linda	104398	10505	0.10	9167	0.09	0	0.00
Baker, Lynne Rudder	282194	15560	0.06	8693	0.03	19832	0.07
Haslanger, Sally	130402	8687	0.07	0	0.00	35968	0.28
Thomasson, Amie	149391	24251	0.16	24900	0.17	17506	0.12
Hurley, Susan	178397	0	0.00	44627	0.25	0	0.00
Longino, Helen	88088	8158	0.09	7383	0.08	0	0.00
MacKinnon, Catharine	17309	0	0.00	0	0.00	0	0.00
Marcus, Ruth Barcan	86796	18713	0.22	0	0.00	0	0.00
Benhabib, Seyla	70781	0	0.00	13467	0.19	0	0.00
Paul, L. A.	122513	33520	0.27	19781	0.16	40369	0.33
Alcoff, Linda Martín	111627	9012	0.08	48302	0.43	8644	0.08

Gendler, Tamar	136347	27803	0.20	28746	0.21	0	0.00
Wolf, Susan	108413	0	0.00	25412	0.23	0	0.00
Adams, Marilyn McCord	151166	34265	0.23	3414	0.02	26616	0.18
Baier, Annette	284507	31533	0.11	17582	0.06	0	0.00
Kamm, Frances	309838	70047	0.23	67616	0.22	62087	0.20
Langton, Rae	121281	41305	0.34	0	0.00	0	0.00
Lloyd, Elisabeth	127173	0	0.00	16064	0.13	43096	0.34

Appendix II. Counts and ratios of argument types by philosopher in the 6-word dataset

Philosopher	**Total**	**De count**	**De ratio**	**In count**	**In ratio**	**Ab count**	**Ab ratio**
Lewis, David K.	295584	49653	0.17	51843	0.18	0	0.00
Quine, W.V.O.	453628	73017	0.16	20889	0.05	0	0.00
Putnam, Hilary	498836	108501	0.22	103907	0.21	40425	0.08
Rawls, John	193515	19669	0.10	82248	0.43	0	0.00
Davidson, Donald	238097	44791	0.19	25158	0.11	4703	0.02
Kripke, Saul	70655	23340	0.33	29528	0.42	0	0.00
Williams, Bernard	229423	80110	0.35	56209	0.25	8728	0.04
Nozick, Robert	84856	2412	0.03	40319	0.48	0	0.00

Williamson, Timothy	574871	149035	0.26	94934	0.17	7862	0.01
Jackson, Frank	456979	59185	0.13	91772	0.20	36922	0.08
Nagel, Thomas	172782	20542	0.12	38077	0.22	29911	0.17
Searle, John R.	247454	28366	0.11	30105	0.12	18833	0.08
Van Fraassen, Bas	492842	122121	0.25	85962	0.17	21250	0.04
Armstrong, David M.	274647	65720	0.24	75368	0.27	9585	0.03
Dummett, Michael	295445	138360	0.47	63435	0.21	9333	0.03
Fodor, Jerry	317155	89400	0.28	113261	0.36	25259	0.08
Harman, Gilbert	348876	45988	0.13	90687	0.26	70541	0.20
Chisholm, Roderick	479455	75549	0.16	32878	0.07	3374	0.01
Dennett, Daniel C.	294155	57321	0.19	55117	0.19	25141	0.09
Chalmers, David J.	213641	120095	0.56	78430	0.37	38191	0.18
Strawson, P.F.	341154	76163	0.22	63110	0.18	12747	0.04
Stalnaker, Robert	389202	40873	0.11	31750	0.08	20989	0.05
Scanlon, T.M.	182519	38246	0.21	61457	0.34	4239	0.02
Dworkin, Ronald	154721	67138	0.43	77536	0.50	9085	0.06
Pettit, Philip	465236	106800	0.23	255681	0.55	8369	0.02
Fine, Kit	620917	206725	0.33	28776	0.05	47326	0.08

Sober, Elliott	528653	73799	0.14	149226	0.28	23369	0.04
Van Inwagen, Peter	455873	110539	0.24	148423	0.33	18321	0.04
Popper, Karl	270095	45727	0.17	35869	0.13	0	0.00
Parfit, Derek	181861	0	0.00	78830	0.43	0	0.00
Kitcher, Philip	538719	43625	0.08	135487	0.25	59603	0.11
Bennett, Jonathan	415384	94822	0.23	107436	0.26	11818	0.03
Nussbaum, Martha	476306	100209	0.21	313152	0.66	0	0.00
Anscombe, G.E.M.	129188	22151	0.17	14230	0.11	0	0.00
Korsgaard, Christine	152127	70678	0.46	15783	0.10	10592	0.07
Anderson, Elizabeth	216245	26240	0.12	39723	0.18	10830	0.05
Thomson, Judith Jarvis	305937	67227	0.22	59026	0.19	0	0.00
Cartwright, Nancy	256543	4962	0.02	72406	0.28	9833	0.04
Annas, Julia	245770	48301	0.20	52757	0.21	15863	0.06
Young, Iris Marion	141579	19189	0.14	84572	0.60	9967	0.07
Millikan, Ruth G.	250734	32194	0.13	102105	0.41	27867	0.11
Foot, Philippa	122136	67347	0.55	13076	0.11	15497	0.13
Stump, Eleonore	200811	65976	0.33	24342	0.12	15717	0.08
Okin, Susan Moller	160664	54285	0.34	114006	0.71	0	0.00

Butler, Judith	22587	935	0.04	0	0.00	0	0.00
O'Neill, Onora	134399	9833	0.07	5792	0.04	0	0.00
Zagzebski, Linda	104398	19672	0.19	29645	0.28	0	0.00
Baker, Lynne Rudder	282194	27321	0.10	8693	0.03	17024	0.06
Haslanger, Sally	130402	8687	0.07	29676	0.23	17984	0.14
Thomasson, Amie	149391	41398	0.28	30645	0.21	24649	0.16
Hurley, Susan	178397	10894	0.06	44627	0.25	0	0.00
Longino, Helen	88088	8158	0.09	21754	0.25	0	0.00
MacKinnon, Catharine	17309	0	0.00	0	0.00	0	0.00
Marcus, Ruth Barcan	86796	22986	0.26	0	0.00	0	0.00
Benhabib, Seyla	70781	0	0.00	13467	0.19	0	0.00
Paul, L. A.	122513	49306	0.40	37201	0.30	30075	0.25
Alcoff, Linda Martín	111627	21342	0.19	57137	0.51	17656	0.16
Gendler, Tamar	136347	27803	0.20	43495	0.32	16662	0.12
Wolf, Susan	108413	9971	0.09	32796	0.30	0	0.00
Adams, Marilyn McCord	151166	47573	0.31	25156	0.17	13308	0.09
Baier, Annette	284507	64105	0.23	37183	0.13	0	0.00
Kamm, Frances	309838	106088	0.34	81853	0.26	42206	0.14

Langton, Rae	121281	66142	0.55	67192	0.55	0	0.00
Lloyd, Elisabeth	127173	5528	0.04	20066	0.16	32173	0.25

Appendix III. Counts and ratios of argument types by philosopher in the 10-word dataset

Philosopher	**Total**	**De count**	**De ratio**	**In count**	**In ratio**	**Ab count**	**Ab ratio**
Lewis, David K.	295584	70514	0.24	82092	0.28	0	0.00
Quine, W.V.O.	453628	88163	0.19	30645	0.07	18129	0.04
Putnam, Hilary	498836	191218	0.38	109692	0.22	40425	0.08
Rawls, John	193515	19669	0.10	83445	0.43	0	0.00
Davidson, Donald	238097	58745	0.25	31495	0.13	16259	0.07
Kripke, Saul	70655	23340	0.33	29528	0.42	0	0.00
Williams, Bernard	229423	87463	0.38	64937	0.28	8728	0.04
Nozick, Robert	84856	19765	0.23	46356	0.55	0	0.00
Williamson, Timothy	574871	192474	0.33	131396	0.23	25389	0.04
Jackson, Frank	456979	87131	0.19	113762	0.25	47338	0.10
Nagel, Thomas	172782	20542	0.12	61894	0.36	36462	0.21
Searle, John R.	247454	50112	0.20	30105	0.12	51120	0.21
Van Fraassen, Bas	492842	157021	0.32	111184	0.23	26494	0.05

Armstrong, David M.	274647	81877	0.30	112514	0.41	19029	0.07
Dummett, Michael	295445	138360	0.47	81554	0.28	9333	0.03
Fodor, Jerry	317155	96104	0.30	157550	0.50	25259	0.08
Harman, Gilbert	348876	62651	0.18	93219	0.27	97364	0.28
Chisholm, Roderick	479455	111919	0.23	43327	0.09	3374	0.01
Dennett, Daniel C.	294155	57321	0.19	73517	0.25	41754	0.14
Chalmers, David J.	213641	120095	0.56	134581	0.63	38191	0.18
Strawson, P.F.	341154	149409	0.44	79438	0.23	12747	0.04
Stalnaker, Robert	389202	57894	0.15	55484	0.14	31039	0.08
Scanlon, T.M.	182519	38246	0.21	61457	0.34	4239	0.02
Dworkin, Ronald	154721	67138	0.43	86621	0.56	32041	0.21
Pettit, Philip	465236	120110	0.26	291073	0.63	8369	0.02
Fine, Kit	620917	249149	0.40	71419	0.12	102826	0.17
Sober, Elliott	528653	87150	0.16	185830	0.35	22873	0.04
Van Inwagen, Peter	455873	158233	0.35	163683	0.36	28842	0.06
Popper, Karl	270095	71080	0.26	43139	0.16	0	0.00
Parfit, Derek	181861	11839	0.07	82200	0.45	25564	0.14
Kitcher, Philip	538719	65778	0.12	203166	0.38	72015	0.13

Bennett, Jonathan	415384	141367	0.34	151836	0.37	11818	0.03
Nussbaum, Martha	476306	122201	0.26	322362	0.68	7043	0.01
Anscombe, G.E.M.	129188	37217	0.29	31063	0.24	0	0.00
Korsgaard, Christine	152127	93817	0.62	15783	0.10	10592	0.07
Anderson, Elizabeth	216245	26240	0.12	73384	0.34	10830	0.05
Thomson, Judith Jarvis	305937	86152	0.28	90458	0.30	0	0.00
Cartwright, Nancy	256543	4962	0.02	93467	0.36	16480	0.06
Annas, Julia	245770	67489	0.27	79609	0.32	15863	0.06
Young, Iris Marion	141579	31211	0.22	108343	0.77	9967	0.07
Millikan, Ruth G.	250734	60788	0.24	136294	0.54	52496	0.21
Foot, Philippa	122136	72504	0.59	21965	0.18	15497	0.13
Stump, Eleonore	200811	84173	0.42	49873	0.25	15717	0.08
Okin, Susan Moller	160664	63453	0.39	155154	0.97	0	0.00
Butler, Judith	22587	935	0.04	0	0.00	0	0.00
O'Neill, Onora	134399	11132	0.08	14611	0.11	9965	0.07
Zagzebski, Linda	104398	30981	0.30	47261	0.45	0	0.00
Baker, Lynne Rudder	282194	27321	0.10	8693	0.03	17024	0.06
Haslanger, Sally	130402	30273	0.23	41222	0.32	26093	0.20

Thomasson, Amie	149391	52082	0.35	41329	0.28	24649	0.16
Hurley, Susan	178397	10894	0.06	81461	0.46	0	0.00
Longino, Helen	88088	14358	0.16	27930	0.32	3810	0.04
MacKinnon, Catharine	17309	0	0.00	17309	1.00	0	0.00
Marcus, Ruth Barcan	86796	22986	0.26	5781	0.07	0	0.00
Benhabib, Seyla	70781	13467	0.19	13467	0.19	0	0.00
Paul, L. A.	122513	49306	0.40	37201	0.30	40551	0.33
Alcoff, Linda Martín	111627	21342	0.19	80137	0.72	17656	0.16
Gendler, Tamar	136347	27803	0.20	43495	0.32	36479	0.27
Wolf, Susan	108413	9971	0.09	32796	0.30	0	0.00
Adams, Marilyn McCord	151166	50971	0.34	25156	0.17	40368	0.27
Baier, Annette	284507	76712	0.27	37183	0.13	0	0.00
Kamm, Frances	309838	120111	0.39	94924	0.31	48836	0.16
Langton, Rae	121281	72420	0.60	50376	0.42	0	0.00
Lloyd, Elisabeth	127173	23647	0.19	47253	0.37	32173	0.25